View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Queensland eSpace

JOURNAL OF BACTERIOLOGY, Aug. 2005, p. 5136–5145 Vol. 187, No. 15
0021-9193/05/$08.00+0 doi:10.1128/JB.187.15.5136–5145.2005
Copyright © 2005, American Society for Microbiology. All Rights Reserved.

Distribution, Expression, and Motif Variability of Ankyrin Domain Genes in *Wolbachia pipientis*†

Iñaki Iturbe-Ormaetxe, Gaelen R. Burke, Markus Riegler, and Scott L. O'Neill*

School of Integrative Biology, The University of Queensland, St. Lucia, QLD 4072, Brisbane, Australia

Received 11 January 2005/Accepted 28 April 2005

The endosymbiotic bacterium *Wolbachia pipientis* infects a wide range of arthropods, in which it induces a variety of reproductive phenotypes, including cytoplasmic incompatibility (CI), parthenogenesis, male killing, and reversal of genetic sex determination. The recent sequencing and annotation of the first *Wolbachia* genome revealed an unusually high number of genes encoding ankyrin domain (ANK) repeats. These ANK genes are likely to be important in mediating the *Wolbachia*-host interaction. In this work we determined the distribution and expression of the different ANK genes found in the sequenced *Wolbachia w*Mel genome in nine *Wolbachia* strains that induce different phenotypic effects in their hosts. A comparison of the ANK genes of *w*Mel and the non-CI-inducing *w*Au *Wolbachia* strain revealed significant differences between the strains. This was reflected in sequence variability in shared genes that could result in alterations in the encoded proteins, such as motif deletions, amino acid insertions, and in some cases disruptions due to insertion of transposable elements and premature stops. In addition, one *w*Mel ANK gene, which is part of an operon, was absent in the *w*Au genome. These variations are likely to affect the affinity, function, and cellular location of the predicted proteins encoded by these genes.

The gram-negative obligate intracellular bacterium *Wolbachia pipientis* is extremely widespread, infecting 20 to 75% of all insect species (21, 59), as well as other invertebrates, such as spiders, mites, terrestrial crustaceans, and filarial nematodes (7, 11, 31, 39, 51). *Wolbachia* is maternally transmitted and can rapidly invade insect populations through the reproductive distortions that it generates in hosts. These include cytoplasmic incompatibility (CI), parthenogenesis, male killing, and reversal of genetic sex determination (33, 38, 47, 48). CI is a type of embryonic lethality that in its simplest form results when *Wolbachia*-infected males mate with uninfected females. CI provides a reproductive advantage to infected hosts and as a result enhances the transmission of *Wolbachia* in host populations.

Despite considerable interest in *Wolbachia* as an agent that might promote insect speciation (5, 10) and as an applied tool for insect pest and disease control (4, 22, 45, 57), little is known about the molecular mechanisms that mediate the various reproductive distortions that it generates. Interestingly, the recent sequencing and annotation of the first *Wolbachia* genome, that of the strain that naturally infects *Drosophila melanogaster* (*w*Mel) (60), revealed an unusually large number of genes that encode proteins containing ankyrin repeat (ANK) domains. While these domains are relatively common in both eukaryotic and viral proteins (42) and have been identified in more than 3,600 different proteins to date (29), they are relatively rare in bacteria. The annotation of 23 ANK genes in the *w*Mel genome (2% of the total number of genes) is very atypical compared to the genomes of related α-*Proteobacteria*, such as *Rickettsia*, *Anaplasma*, and *Ehrlichia*, whose genomes typically contain only one to three genes encoding ankyrin repeats (2, 12). It has been proposed that the ANK genes in *Wolbachia* are likely to play a functional role in its unique biology (60).

The ANK domain is typically a 33-residue L-shaped motif containing two antiparallel α-helices connected by a short loop. ANK domains mediate protein-protein interactions (29, 43) in diverse families of proteins, including cytoskeletal and membrane proteins, transcriptional and developmental regulators, toxins, and CDK (cyclin-dependent kinase) inhibitors (6, 27, 43). Interestingly, the inhibition of CDK1 has been proposed as a possible mechanism explaining the CI phenotype induced by *Wolbachia* in *Nasonia* wasps (53, 54). Moreover, in the related intracellular tick-borne pathogen *Anaplasma phagocytophilum* an ANK protein (AnkA) is secreted into the host cell, where it binds host chromatin, suggesting that it has a role in the regulation of host gene expression (12).

Considering the potential importance of ANK motifs in mediating protein-protein interactions and their profusion in *Wolbachia*, we performed a comparative study to examine the distribution, transcription, and sequence variation of ANK genes from nine different *Wolbachia* strains (Table 1). These strains all infect *Drosophila* and are capable of generating a range of different CI crossing types, and in some cases they are unable to cause CI (17). The latter strains are known to be incapable of inducing CI in males but still retain the capacity to rescue CI in females ($mod^-/resc^+$) (8, 26), or they are incapable of either inducing or rescuing CI ($mod^-/resc^-$).

* Corresponding author. Mailing address: School of Integrative Biology, The University of Queensland, St. Lucia, QLD 4072, Brisbane, Australia. Phone: 61 7 33652471. Fax: 61 7 33469213. E-mail: scott.oneill@uq.edu.au.

† Supplemental material for this article may be found at http://jb.asm.org/.

MATERIALS AND METHODS

Fly stocks and *Wolbachia* strains. The *Wolbachia* and *Drosophila* strains used in this work are listed in Table 1. *Wolbachia* strains were selected on the basis of the extent to which they cause CI (strong, weak, or non-CI inducers) in their hosts and their modification/rescue phenotypes. *Wolbachia* infections were main-

Downloaded from http://jb.asm.org/ on October 13, 2015 by University of Queensland Library

TABLE 1. *Wolbachia* strains and *Drosophila* hosts used in this work

Wolbachia strain	*Drosophila* host	Phenotype	Reference(s)[a]
*w*Mel	*D. melanogaster* yw67c23	mod$^+$/resc$^+$	16, 46
*w*MelPop	*D. melanogaster* W1118	mod$^+$/resc$^+$	24, 28
*w*MelCS	*D. melanogaster* CS	mod$^+$/resc$^+$	18, 46
*w*Cer2	*D. simulans* RC21[b]	mod$^+$/resc$^+$	36
*w*Au^1	*D. simulans* Coffs Harbour	mod$^-$/resc$^-$	62
*w*Au^2	*D. simulans* Coffs Harbour	mod$^-$/resc$^-$	35
*w*Au^3	*D. simulans* Coffs Harbour	mod$^-$/resc$^-$	17
*w*Au^4	*D. simulans* Y6 Yaounde	mod$^-$/resc$^-$	13
*w*Ri	*D. simulans* DSR	mod$^+$/resc$^+$	19
*w*Ha	*D. simulans* DSH	mod$^+$/resc$^+$	34
*w*No	*D. simulans* N7NO	mod$^+$/resc$^+$	25
*w*Mau	*D. simulans* DSW(Ma)	mod$^-$/resc$^+$	15, 37

[a] Reference(s) in which the strain and/or the phenotype was first described.
[b] Transinfected from *Rhagoletis cerasi*.

tained in fly stocks reared on standard corn flour-sugar-yeast medium at 25°C. Clearing of *Wolbachia* infections with tetracycline was performed as described elsewhere (18).

Dot blot analysis. DNA was extracted from *D. melanogaster* or *Drosophila simulans* female flies harboring the different *Wolbachia* strains (Table 1). Flies were homogenized and extracted by using either the Holmes-Bonner protocol (20) or an STE extraction method (32). DNA from tetracycline-treated *D. melanogaster* yw^{67c23} (*w*Mel-T) and *D. simulans* Riverside-DSR (*w*Ri-T) flies was also extracted and used as negative controls. DNA was spotted onto Zeta-Probe nylon filters (Bio-Rad), cross-linked by UV irradiation, and hybridized at 65°C overnight in 0.5 M sodium phosphate buffer, pH 7.0, 7% sodium dodecyl sulfate, 1 mM EDTA. Following hybridization the membranes were washed under medium-stringency conditions at 65°C (15-min washes in 2× SSC, 1× SSC, and then 0.5× SSC, all containing 0.1% sodium dodecyl sulfate [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate]). Autoradiography was performed with a PhosphorImager screen (Molecular Dynamics).

The probes for the 23 different ANK genes (see Fig. S1 in the supplemental material) were PCR amplified from *Wolbachia w*Mel genomic DNA using specific primers (see Table S1 in the supplemental material). The *Wolbachia* surface protein gene *wsp* was amplified with primers 81F and 691R (9) and probed as a control for *Wolbachia* DNA. The PCR cycling conditions were as follows: 94°C for 3 min, followed by 94°C for 30 s, 50°C for 30 s, and 72°C for 3 min for 35 cycles and then 72°C for 10 min. The reaction mixture (final volume, 20 μl) contained each primer at a concentration of 500 nM, each deoxynucleoside triphosphate at a concentration of 200 μM, 1.5 mM $MgCl_2$, 100 ng of *w*Mel DNA, and 1 U of *Taq* polymerase (Promega). The reaction buffer contained 10 mM Tris (pH 9.0), 50 mM KCl, and 0.1% Triton X-100. PCR products were separated in 1% agarose gels, stained with ethidium bromide, and gel purified using gel extraction kits (QIAGEN). They were radioactively labeled with [α-^{32}P]dATP (Amersham Pharmacia) using a Random Primed DNA labeling kit (Roche) and were cleaned prior to hybridization with a PCR purification kit (QIAGEN).

RT-PCR. Total RNA was extracted from *Drosophila* harboring the different *Wolbachia* strains using Trizol (Invitrogen), followed by chloroform extraction and isopropanol precipitation. The RNA preparation was treated with RNase-free RQ1 DNase (Promega), and first-strand cDNA was synthesized from 5 μg of total RNA using reverse transcriptase (RT) (Superscript III; Invitrogen) and random primers (Promega) at 42°C for 60 min. The cDNA was treated with RNase H prior to the PCR. Negative controls to detect genomic DNA contamination were processed in the same way, except that no reverse transcriptase was added to the reaction mixture. cDNA synthesized using RNA from tetracycline-treated *Drosophila* flies was used as a *Wolbachia*-free negative control. PCR amplification was performed as described previously using 1 μl of cDNA as the template and the primers listed in Table S1 in the supplemental material. Each RT-PCR was repeated three times using independent RNA extracts and cDNA synthesis reactions. Negative controls showed no PCR amplification. In order to characterize expression of the WD0512-WD0513-WD0514 operon (see Fig. 2B), the following primers spanning the intergenic regions were used: P1 (5′-CTAA TGCAAACCCATGAAACCCTGC-3′), P2 (5′-CCATTTATAATAGCTGGG GCTATGG-3′), P3 (5′-GAGAATTATCTTGATAGAGTTGTACC-3′), and P4 (5′-CGATATTGTTTTAGAGAAAACAAAGG-3′).

Characterization of the genomic region around WD0514 in the *Wolbachia w*Au strain. For sequencing of the genomic region flanking WD0514, DNA was extracted from the *Wolbachia w*Au strain, digested using either the EcoRI or SpeI restriction endonuclease (New England Biolabs), and ligated overnight at 12°C into pBluescript II SK (Stratagene). The ligation mixtures were diluted 1:20, and 1 μl was used as a template for PCRs performed with various forward primers specific for several open reading frames (ORFs) adjacent to WD0514 and reverse primers specific for pBluescript, such as primer M13R (5′-CAGGA AACAGCTATGAC-3′) or T7 (5′-TAATACGACTCACTATAGGG-3′). The PCR conditions were the same as those described above for the RT-PCR analysis, except that Expand High Fidelity *Taq* polymerase (Roche) was used. PCR bands were cloned into the pGEM-T Easy vector (Promega) and sequenced with the T7 and M13R universal primers using an AB Big Dye terminator kit (version 3.1) with fluorescent sequencing and AmpliTaq DNA polymerase (Perkin-Elmer), and they were analyzed with an AB 3730xl-96 capillary sequencer. Sequencing was done at the Australian Genome Research Facility. Sequence similarity searches were performed using the BLAST algorithm (1) at the National Center for Biotechnology Information. Analysis and assembly of the sequences were done using the *EditSeq*, *SeqMan*, and *MegAlign* components of the Lasergene sequence analysis software package (DNAStar Inc., Madison, Wis.). The primers used to confirm the insertion size (see Fig. 3B) were P5 (5′-GCA GCCATGCTCGGTAA-3′) and P6 (5′-ACTTTGGAGTTAAAACCGTA-3′). These primers are 28.25 kb apart in the *w*Mel genome and anneal to single-copy genes (WD0505 and WD0523, respectively).

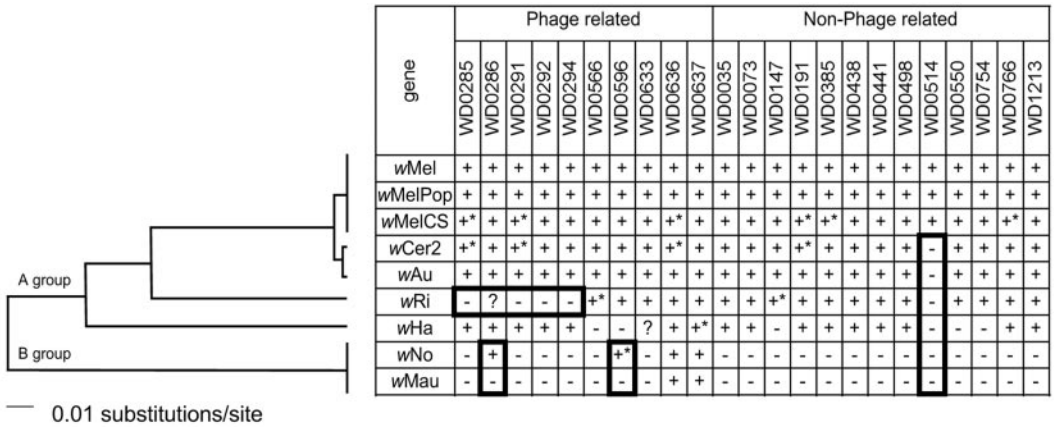
Sequencing of ANK genes in the *Wolbachia w*Au strain. ANK genes from the different *Wolbachia* strains were PCR amplified using specific primers (see Table S1 in the supplemental material) based on the recently sequenced *Wolbachia* strain *w*Mel genome (60). The PCR parameters were basically those described above, except that Expand High Fidelity *Taq* polymerase (Roche) was used. Three or four independent PCRs were performed for each gene. Sequence manipulation and analysis were done as described above. The protein domains were identified by using SMART v3.5 (http://smart.embl-heidelberg.de/) (23, 42). The partial sequences which we obtained for the WD0191 and WD0285 genes in *w*Au using the primers indicated in Table S1 are identical to the sequences of the genes in *w*Mel.

Nucleotide sequence accession numbers. Partial sequences of the *w*Au genes that differ from their homologues in *w*Mel have been deposited in the GenBank database under the following accession numbers: WD0035, AY971752; WD0073, AY971753; WD0147, AY971754; WD0286, AY971755; WD0291, AY971756; WD0292, AY649749; WD0294, AY649750; WD0385, AY664873; WD0438, AY971757; WD0441, AY971758; WD0498, AY836559; WD0550, AY649751; WD0566, AY971759; WD0596, AY971760; WD0633, AY672910; WD0636, AY649752; WD0637, AY971761; WD0754, AY836560; WD0766, AY649753; and WD1213, AY971762.

RESULTS

Distribution of ANK genes in different *Wolbachia* strains. We determined the distribution of the 23 ANK genes found in the *w*Mel *Wolbachia* genome in a variety of *Wolbachia* strains that have different effects in their hosts (Table 1). We placed special emphasis on the comparison between the sequenced *w*Mel strain, which causes CI in *Drosophila*, and the closely related *w*Au strain. The latter strain was initially found in Australia and America (17, 56) and is unable to cause CI in *Drosophila*. Probes for the 23 ANK genes (see Fig. S1 in the supplemental material) were PCR amplified, radioactively labeled, and hybridized to membranes containing DNA from a wide range of *Wolbachia* strains. We found that the distribution of the 23 *w*Mel genes coding for ANK proteins varied significantly in the strains (Fig. 1). A dot blot and PCR analysis was especially useful for revealing putatively absent genes. As expected, all the probes hybridized to *w*Mel DNA and to DNA of the virulent *w*MelPop ("popcorn") strain, whose genome sequence is very similar to that of *w*Mel (49). Only 2 of the 23 ANK genes (WD0441 and WD0498) weakly hybridized (data not shown) to DNA extracted from the tetracycline-treated

Downloaded from http://jb.asm.org/ on October 13, 2015 by University of Queensland Library



gene	Phage related									Non-Phage related												
	WD0285	WD0286	WD0291	WD0292	WD0294	WD0566	WD0596	WD0633	WD0636	WD0637	WD0035	WD0073	WD0147	WD0191	WD0385	WD0438	WD0441	WD0498	WD0514	WD0550	WD0754	WD0766	WD1213
*w*Mel	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
*w*MelPop	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
*w*MelCS	+*	+	+*	+	+	+	+	+*	+	+	+	+	+	+*	+*	+	+	+	+	+	+	+*	+
*w*Cer2	+*	+	+*	+	+	+	+	+*	+	+	+	+	+	+*	+	+	+	+	-	+	+	+	+
*w*Au	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+
*w*Ri	-	?	-	-	-	+*	+	+	+	+	+	+	+*	+	+	+	+	+	-	+	+	+	+
*w*Ha	+	+	+	+	+	-	-	?	+	+*	+	+	-	+	+	+	+	+	-	-	-	+	+
*w*No	-	+	-	-	-	-	+*	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
*w*Mau	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-

FIG. 1. Distribution of ANK genes in several *Wolbachia* strains. The plus and minus signs indicate the presence and absence, respectively, of an above-background signal in dot blot hybridization experiments. An asterisk indicates that data for samples were confirmed by PCR. The dot blot results for the four *w*Au strains used in this study (Table 1) were identical, and they are grouped for clarity. The most interesting results are enclosed in boxes. The phylogenetic tree on the left is a cladogram based on the *Wolbachia* surface protein gene (*wsp*) sequences of strains *w*Mel (GenBank accession number AF020072), *w*MelPop (AF338346), *w*MelCS (AF020065), *w*Cer2 (AF418557), *w*Au (AF020067), *w*Ri (AF20070), *w*Ha (AF020068), *w*No (AF020074), and *w*Mau (AF020069); this cladogram was constructed by the unweighted-pair group method using average linkages. The four *w*Au strains used in this study have identical *wsp* sequences. Sequences were aligned using Clustal X (52), and the phylogenetic tree was constructed using PAUP (50).

flies (*w*Mel-T and *w*Ri-T), which were used as negative controls. This was most likely the result of nonspecific hybridization, as DNA from these flies was negative (as determined both by PCR and by hybridization) for the characteristic *Wolbachia* *wsp* gene (Fig. 2A, bottom panel). In this work we used four *w*Au strains ($mod^-/resc^-$) obtained from different fly stocks (Table 1) in order to compare the reproducibility of the results and the genetic consistency of the infection. The four strains

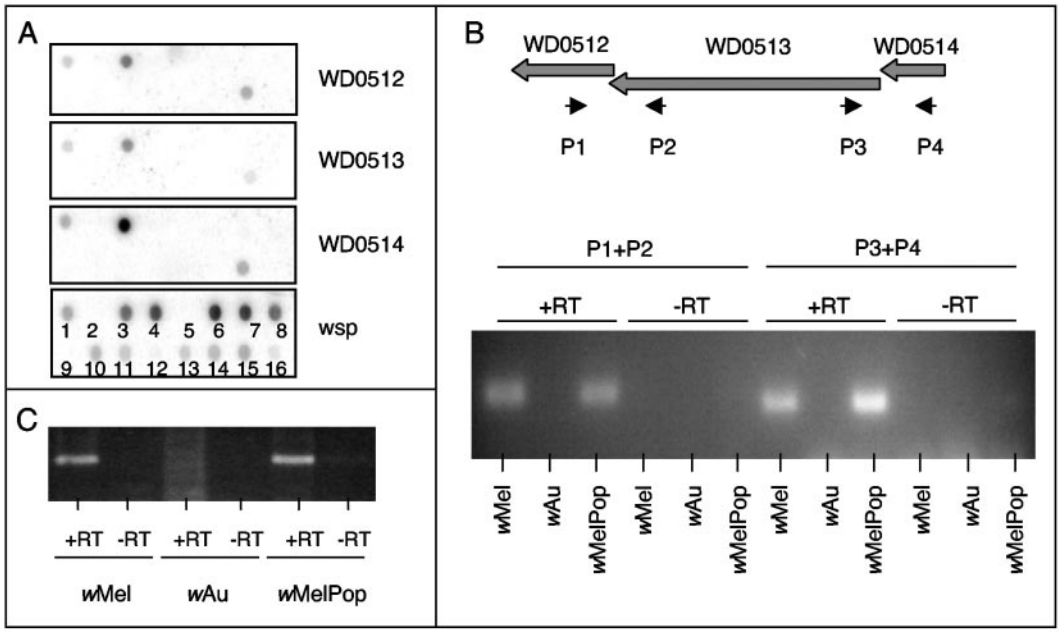


FIG. 2. WD0512-WD0513-WD0514 is an operon that is not present in the *Wolbachia* *w*Au strain. (A) Dot blot hybridization analysis of the WD0512, WD0513, and WD0514 genes in all the *Wolbachia* strains used in this study. A control dot blot for the *wsp* gene was also included. Spot 1, *w*Mel; spot 2, *w*Mel-T, tetracycline treated; spot 3, *w*MelPop; spot 4, *w*Ri; spot 5, *w*Ri-T, tetracycline treated; spot 6, wAu^1; spot 7, wAu^2; spot 8, wAu^3; spot 9, negative control; spot 10, wAu^4; spot 11, *w*No; spot 12, *w*Mau; spot 13, *w*Cer2; spot 14, wAu^4; spot 15, *w*MelCS; spot 16, *w*Ha. Spot 9 contained extraction buffer and was used as a negative control. Spot 14 was a duplicate of spot 10 and was used as a reproducibility control. (B) RT-PCR demonstrating the expression of the three genes as a single transcript in the *w*Mel and *w*MelPop *Wolbachia* strains. Lanes 1 to 3, primers spanning the junction between WD0512 and WD0513 (P1+P2); lanes 4 to 6, negative controls (−RT); lanes 7 to 9, primers spanning the junction between WD0513 and WD0514 (P3+P4); lanes 10 to 12, negative controls (−RT). (C) RT-PCR of WD0514 in the *w*Mel, *w*Au, and *w*MelPop *Wolbachia* strains, showing no expression of this gene in *w*Au. Negative controls, in which no reverse transcriptase was added during the cDNA synthesis (−RT), were included.

Downloaded from http://jb.asm.org/ on October 13, 2015 by University of Queensland Library

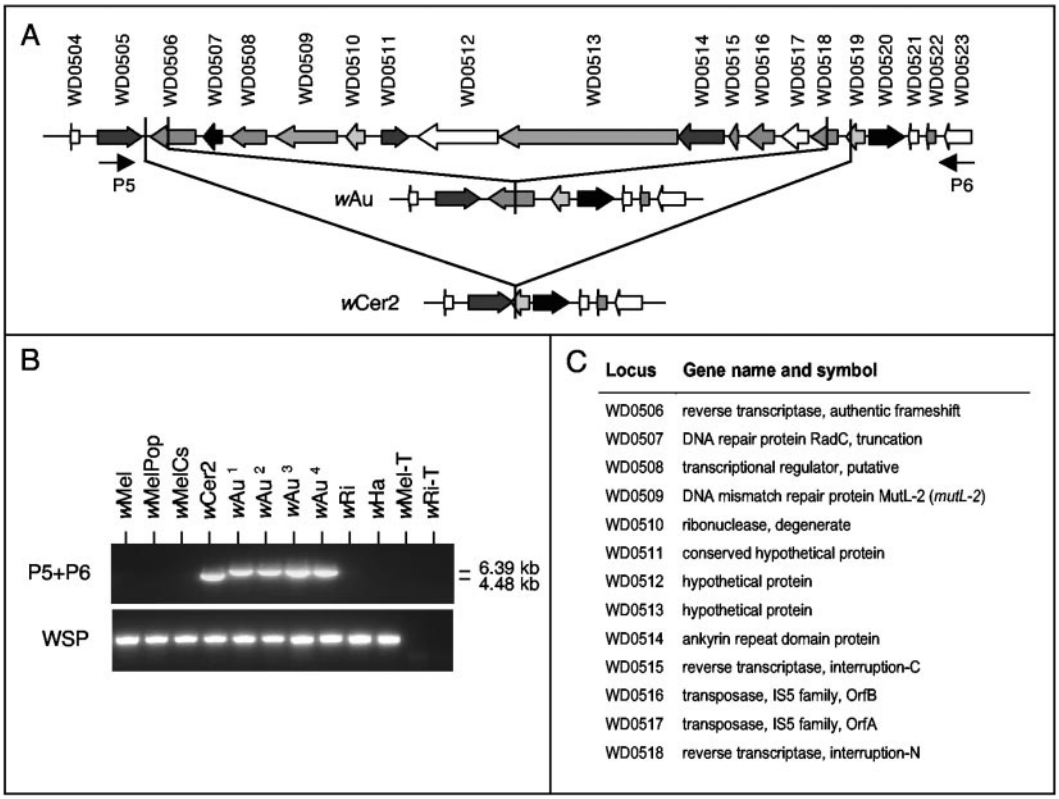


FIG. 3. (A) *w*Mel genomic region surrounding the WD0514 ANK gene, showing the genes that are missing in the *w*Au and *w*Cer2 strains. (B) (Top panel) PCR for the region in various *Wolbachia* strains using primers P5 and P6 specific for the single-copy genes WD0505 and WD0523, which are 28.25 kb apart in the *w*Mel genome. The PCR showed that the missing fragment in *w*Au is 21.86 kb long (28.25 kb minus 6.39 kb), and in *w*Cer2 is 23.77 kb long (28.25 kb minus 4.48 kb). (Bottom panel) Positive PCR control for *Wolbachia* DNA using primers for the *wsp* gene. Tetracycline-treated flies (*w*Mel-T and *w*Ri-T) were cleared of *Wolbachia*. (C) Genes surrounding the WD0512-WD0513-WD0514 operon in the *w*Mel genome. The annotation and sequence of the genes shown can be found at the Comprehensive Microbial Resource at www.tigr.org.

gave identical results in all experiments. Most ANK genes were found in all group A *Wolbachia*, with a few exceptions. WD0514 was present only in strains that have the *w*Mel CI crossing type (*w*Mel, *w*MelCS, and *w*MelPop), and it was absent in all strains that have different CI crossing types, as well as *w*Au (mod$^-$/resc$^-$) and the *w*Mau strain (mod$^-$/resc$^+$). In addition, a group of phage-associated ANK-containing genes (WD0285, WD0291, WD0292, and WD0294) were absent from *w*Ri (Fig. 1), most likely reflecting differences in prophage insertions between strains. The *w*Ha *Wolbachia* strain, a more distant relative of the *w*Mel strain, as shown in the cladogram in Fig. 1, was negative for 7 of the 23 ANK genes. Most ANK genes were not detected in group B *Wolbachia*, probably due to high levels of sequence divergence between strains since BLAST analysis of the unfinished *Wolbachia* genome of *Culex* (www.sanger.ac.uk) and the recently completed *Wolbachia* genome of *Brugia* (14) revealed a number of genes with similarity to the *w*Mel ANK genes. Comparison of the mod$^-$/resc$^+$ strain *w*Mau with the very closely related mod$^+$/resc$^+$ strain *w*No revealed the absence of WD0286 and WD0596 in *w*Mau (Fig. 1).

ANK gene WD0514 is part of an operon that is present in all mod$^+$ members of the *w*Mel clade and absent in all mod$^-$ *w*Au strains. Since the WD0514 gene was present only in strains belonging to the *w*Mel clade, which were capable of generating the *w*Mel CI crossing type, we examined the possibility that this was the result of recent introduction into this clade. Both dot blot analysis (Fig. 2A) and PCR analysis (data not shown) of the chromosomal region around this gene showed that the upstream ORFs WD0512 and WD0513 were absent in all strains that lacked WD0514, including *w*Au, and were present only in *w*Mel, *w*MelCS, and *w*MelPop. RT-PCR using primers spanning the junction between these three ORFs demonstrated that these three genes, which have little or no intergenic space between them, are transcribed as a single transcriptional unit in *Wolbachia* strains *w*Mel and *w*MelPop (Fig. 2B), whereas there is no expression of the operon or WD0514 in *w*Au (Fig. 2C).

To determine the extent of the presumed insertion around the WD0512-WD0513-WD0514 operon in *w*Mel strains, *w*Au genomic DNA was digested with either EcoRI or SpeI endonuclease, ligated into pBluescript, and used for chromosome walking by PCR performed with primers in the vector and in various single-copy genes contiguous to these ORFs. After sequencing and assembling the resulting PCR products, we identified a difference of 21.86 kb between the *w*Mel genome and the *w*Au strain genome (Fig. 3A). This result was confirmed by PCR amplification of the sequence between the single-copy WD0505 and WD0523 genes that flank this region. The distance between the PCR primers (P5 and P6) (Fig. 3A)

Downloaded from http://jb.asm.org/ on October 13, 2015 by University of Queensland Library

in *w*Mel is 28.25 kb, but the PCR product obtained with *w*Au DNA was 6.39 kb long (Fig. 3B), indicating that there was a 21.86-kb difference compared to *w*Mel strains. Sequencing and restriction analysis of this band confirmed the gap between *w*Mel chromosomal positions 486988 and 508845 that includes the ORFs WD0506 to WD0518. In *w*Cer2, the PCR band obtained was 4.48 kb (Fig. 3B), indicating a larger, 23.77-kb difference (positions 486531 to 510304) compared to the *w*Mel genome. The additional chromosomal section in *w*Mel contains a series of genes (Fig. 3C) related to mobile element function, such as genes encoding two reverse transcriptases that are partially deleted in *w*Au (WD0506 and WD0518) or one reverse transcriptase in *w*Cer2, as well as genes encoding two IS*5* transposases (WD0516 and WD0517) (60). The fragment also contains genes encoding a degenerate RNase, a conserved hypothetical protein, and a transcriptional regulator, as well as two DNA repair genes (*radC* and *mutL-2*). The absence of these genes in *w*Au and *w*Cer2 might have no phenotypic effect if they are multicopy genes, as they are in *w*Mel (there is an extra copy of *radC* plus two truncated copies, and there is a paralogue of *mutL-2*, designated *mutL-1*, as well as a related gene, *mutS*). Of the 13 genes in *w*Mel, only *radC*, *mutL-2*, and the RNase gene have orthologues in both *Ehrlichia* and *Anaplasma*, whereas the rest of the genes, including the WD0512-WD0513-WD0514 operon, have no orthologues in the genomes of these relatives.

ANK proteins are highly variable in mod$^+$/resc$^+$ and mod$^-$/resc$^-$ *Wolbachia* strains. In order to determine possible differences between ANK genes of mod$^+$/resc$^+$ and mod$^-$/resc$^-$ *Wolbachia* strains, we partially sequenced all 22 ANK genes from the *w*Au strain and compared them to the genes of *w*Mel. Thirteen of the 22 genes had minor sequence variations (98 to 100% identity; see Materials and Methods for accession numbers). Among the rest, WD0292 encodes a protein containing a 4-amino-acid insertion in *w*Au (accession no. AY649749) compared to the *w*Mel protein, whereas the protein encoded by WD0633 (accession no. AY672910) in *w*Au has small insertions and deletions of amino acids (two insertions and three deletions). Notably, seven ANK genes in *w*Au encode proteins with important differences (Fig. 4) compared with their *w*Mel homologues, including variations in the number of ANK repeats, ORF disruption by transposable element insertions, premature stops, and fusion to an adjacent ORF, that were initially annotated as separated genes in the *w*Mel genome. As a result of a 66-amino-acid deletion affecting repeats 4 and 5, the phage-associated WD0294 protein (accession no. AY649750) has seven ANK repeats in *w*Au compared with nine ANK repeats in *w*Mel. WD0550 (accession no. AY649751) contains two extra ANK repeats (coding 66 extra amino acids) in *w*Au, whereas WD0766 (accession no. AY649753) contains three extra ANK repeats as a result of two insertions coding for 66 and 58 amino acids. In this case the precise ANK motifs that are deleted are unclear, although it appears that domains 2, 3, and 6 from *w*Au are absent in *w*Mel (Fig. 4). Most importantly, in the *w*Au strain this gene contains a premature stop that eliminates the two transmembrane domains at the C terminus (Fig. 4) of the protein.

We also found that some ANK genes are larger in *w*Au than in *w*Mel, as a result of mutations in the sequences that eliminate the stop codons that separate them from the next gene. On the one hand, WD0498 and WD0499 were initially annotated as separated ORFs in the *w*Mel genome (60), but we found no stop codon between them in the *w*Au strains. The ANK protein encoded by WD0498 is therefore larger in *w*Au (accession no. AY836559) and contains an extra ANK repeat, previously unidentified in the annotation of the *w*Mel genome, as a result of its fusion with WD0499 (Fig. 4). WD0499 is, however, shorter in *w*Au as a result of a premature stop codon. It seems clear that the mutation of a CAA codon into a TAA stop codon that results in the removal of one ANK repeat in the protein occurred in the *w*Mel lineage, since this stop codon also appears in *w*MelCS and *w*MelPop (sequences identical to *w*Mel) but not in *w*Au (accession no. AY836559), *w*Cer2, or *w*Ri (accession no. AY971763; *w*Cer2 sequence similar to the *w*Ri sequence). Similarly, we also found that the stop codon that separates the ANK gene WD0754 from the hypothetical gene WD0753 is present only in *w*Mel, *w*MelCS, and *w*MelPop (similar sequences). In *w*Au (accession no. AY836560) and *w*Cer2 (same sequence) a change from TAA to TCA results in fusion of the two genes as a single coding sequence. In this case, the addition of WD0753 to WD0754 does not modify the number of ANK repeats in the encoded *w*Au protein, but it does add an extra two transmembrane domains, as determined by TMHMM (Fig. 4). Gene junctions between other ANK genes that appear to be immediately adjacent to their flanking genes were also sequenced in various strains, but no other mutations in stop codons were found.

In addition to variation in the number of ANK domains, some ANK genes in the *w*Au *Wolbachia* strain contain major disruptions. The WO phage-associated ANK gene WD0636 (accession no. AY649752) carries a point mutation that introduces a premature stop codon into this gene in all the mod$^-$/resc$^-$ *w*Au strains examined. This mutation is predicted to result in the production of a truncated WD0636 protein that lacks one ANK motif at the carboxy terminus and could affect its function. This premature stop was not found in other *Wolbachia* strains, such as *w*MelPop and *w*MelCS (data not shown). The phage-associated ANK gene WD0385 from *w*Au (accession no. AY664873) was found to contain a full-length 919-bp IS*5* insertion element disrupting the ORF at nucleotide position 769, in the middle of the seventh ANK motif (Fig. 4). IS*5* elements are very common in *w*Mel, and there are 13 identical copies in the chromosome (60). They contain two ORFs for transposases and are flanked by terminal inverted repeats (TIRs). Interestingly, whereas all 13 IS*5* elements in the *w*Mel genome have asymmetrical TIRs containing one mismatch, the IS*5* element inserted into WD0385 in *w*Au is flanked by identical TIRs (5′-AGAGGTTGTCCGGAAACA AGTAAA-3′). The *orfA* gene in this *w*Au IS*5* element encodes a transposase with four amino acid differences compared with the *w*Mel OrfA.

ANK gene expression. The expression of the 23 ANK genes from *w*Mel was determined by RT-PCR by using RNA isolated from the *w*Mel, *w*MelPop, and *w*Au *Wolbachia* strains. Two RT-PCR examples are shown in Fig. 5A. RT-PCRs showed that most ANK genes are actively expressed in these strains; the only exceptions are WD0514 (Fig. 2C), which is missing in the *w*Au chromosome (Fig. 1), and WD0385, which is only partially transcribed in *w*Au (Fig. 5B). When we used RT-PCR primers spanning the junction across the insertion element, we

Downloaded from http://jb.asm.org/ on October 13, 2015 by University of Queensland Library

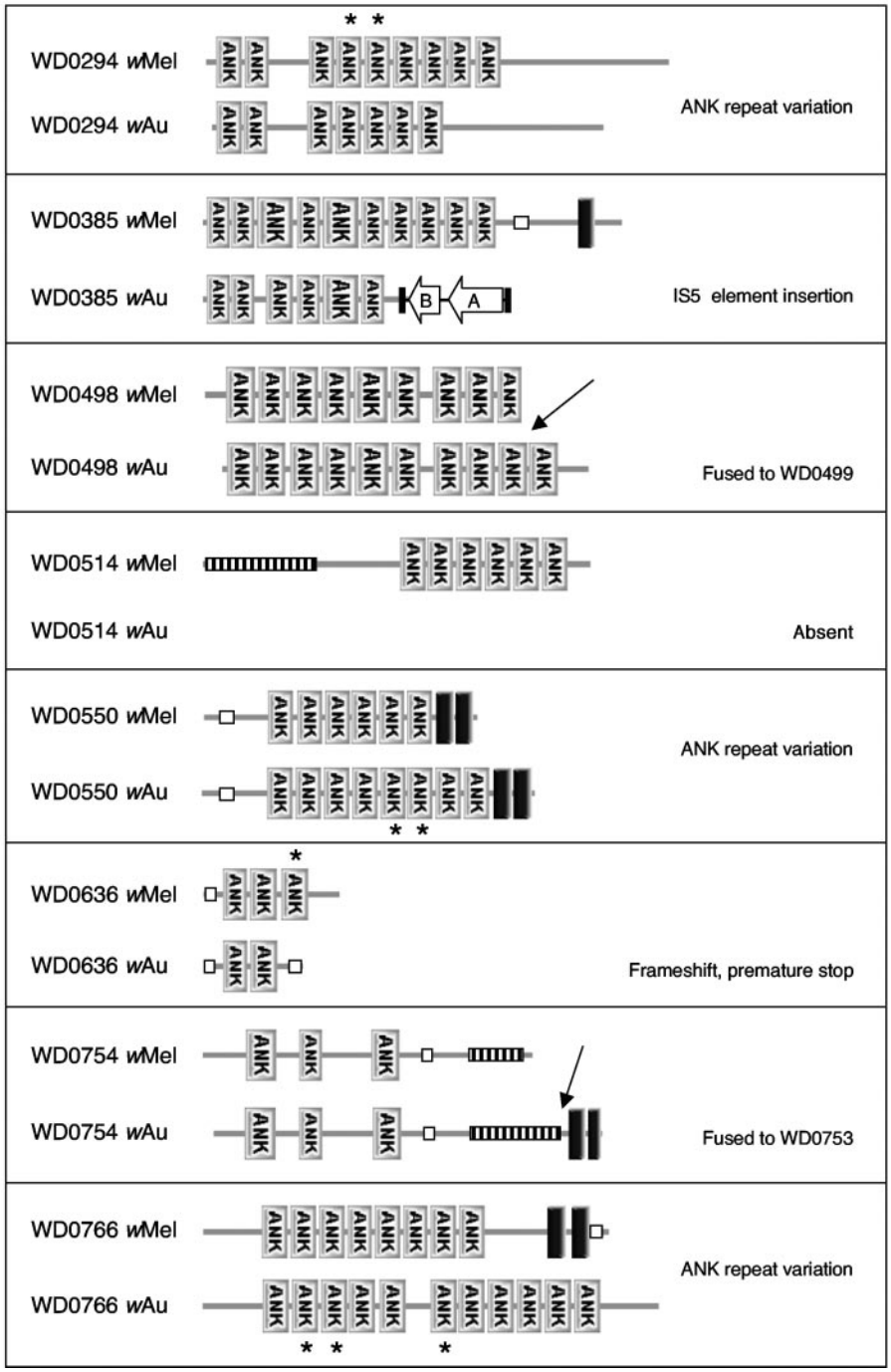


FIG. 4. ANK proteins whose sequence and/or domain architecture is significantly different in *w*Mel and *w*Au. The location of ANK motifs was determined using SMART v3.5 (http://smart.embl-heidelberg.de/) (23, 42). Similar results were obtained by analysis of the Pfam Conserved Domain at Washington University (http://pfam.wustl.edu). Transmembrane domains are represented by black boxes, as predicted by the *TMHMM2* server. Coiled coil regions (striped boxes) were determined by the *Coils2* program. The open squares represent segments with low compositional complexity, as determined by the *SEG* program. The arrows indicate point mutations in the stop codon between WD0498 and WD0499 and between WD0754 and WD0753. Inserted or deleted ANK domains are indicated by asterisks.

found expression of the WD0385 cDNA upstream of the IS*5* insertion element (primers P7 [5′-GCAGAAGATGAAGAG GGAAAC-3′] and P8 [GAGTTCGTATGTCTTGAGTAG]) but not across the insertion point (primers P7 and P9 [5′-AA GGGAATGGTCAAGAATAG-3′]). Early termination of the transcript is probably caused by the TIRs that flank IS*5* and that could act as a transcriptional terminator element by forming a 24-bp hairpin, as determined using secondary RNA prediction programs (http://www.genebee.msu.su). The formation of this hairpin is facilitated by the fact that the TIRs that flank

Downloaded from http://jb.asm.org/ on October 13, 2015 by University of Queensland Library

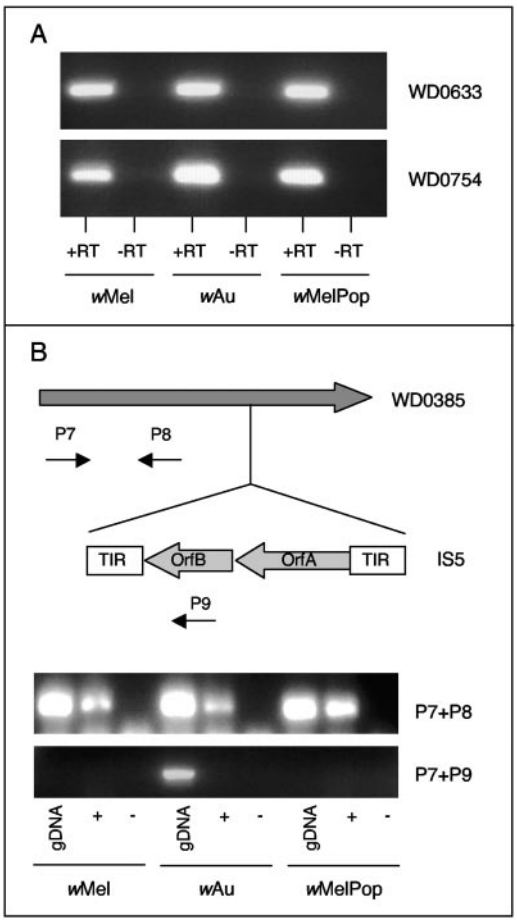


FIG. 5. (A) RT-PCR showing expression of WD0633 and WD0754 in *Wolbachia* strains *w*Mel, *w*Au, and *w*MelPop. −RT, negative controls with no reverse transcriptase. (B) RT-PCR demonstrating partial expression of WD0385 in *Wolbachia* strain *w*Au. This gene contains an IS*5* insertion element, and only primers P7 and P8, upstream the insertion, were able to amplify the cDNA, whereas primers P7 and P9 could amplify only genomic DNA (gDNA) from *w*Au.

this IS*5* element in the WD0385 gene are identical in their 24-bp sequences, unlike the TIRs in all 13 IS*5* elements in the *w*Mel genome, which contain a T/A mismatch at position 18.

DISCUSSION

The sequencing of the *Wolbachia* genome has revealed a surprisingly high number of genes coding for ANK repeats. The fact that these motifs are typically involved in protein-protein interactions makes the proteins very attractive candidates for molecules that are involved in the molecular communication that takes place between *Wolbachia* and the host cell, a process that has not been characterized yet. Analysis of the 23 *w*Mel ANK genes in other group A and B *Wolbachia* strains revealed significant variation between strains. Obviously, the possibility that novel ANK genes not present in the *w*Mel sequence are present in other strains cannot be excluded. In fact, Salzberg et al. (41) recently assembled a nearly complete *Wolbachia* genome from *D. simulans* that contains seven new ANK genes not found in the *w*Mel genome.

Strains that are capable of both inducing the CI modification in *Drosophila* sperm and rescuing this modification in *Drosophila* eggs during fertilization are known as mod$^+$/resc$^+$ strains and include *w*Mel, *w*MelCS, and *w*MelPop (58). In contrast, *w*Au strains, while very closely related, have been shown to be incapable of either modifying host sperm or rescuing the modification of related strains and are designated mod$^-$/resc$^-$ strains. The *w*Mel, *w*MelCS, and *w*MelPop strains and various *w*Au *Wolbachia* strains were compared to examine differences in either the distribution, expression, or sequence of ANK genes that might correlate with the phenotypic differences known to occur in these strains. The first obvious difference observed between strains is the absence of the ankyrin domain gene WD0514 in all *w*Au strains (Fig. 1). When the region surrounding WD0514 was examined in more detail, we found that this gene is part of an operon together with the WD0512 and WD0513 genes. Moreover, the operon was found to be part of a 21.86-kb insertion in the *w*Mel clade that contains 13 ORFs (Fig. 3). Considering that the WD0512-WD0513-WD0514 operon is found only in the *w*Mel clade strains examined in this study, the most parsimonious explanation for the presence of this cluster of genes is an initial insertion of a reverse transcriptase in *w*Au (Fig. 3A), followed by insertion of the 21.86-kb element in the *w*Mel-*w*MelCS-*w*MelPop lineage. This event would have been combined with duplication of the reverse transcriptase in *w*Mel (Fig. 3A) to give the reverse transcriptase genes WD0506 and WD0518. Due to sequence divergence, we could not characterize by PCR this region in other *Wolbachia* strains that also lack the operon, although we found that the region is slightly different in *w*Au and *w*Cer2. It is noteworthy that the G+C content of these 13 genes (35.7%) is similar to the average G+C content (35.2%) of the *Wolbachia w*Mel genome (60), suggesting that the genes either were laterally transferred into the *w*Mel lineage from a donor with a similar G+C content or, more probably, were present in the *Wolbachia* lineage for a considerable amount of time. In fact, the codon usage index in WD0506 to WD0518 is similar to the overall codon usage in the *Wolbachia w*Mel genome (http://www.evolvingcode.net/codon) (44), and this supports the finding that most of the variation in codon bias in the *Wolbachia* genome can be traced to variation in G and C (60). When genes in this region, other than those associated with mobile genetic elements, are compared to the genomes of the related bacteria *Ehrlichia*, *Anaplasma*, and *Rickettsia*, only the operon containing WD0512-WD0513-WD0514 is unique to the *Wolbachia* lineage and exhibits no significant similarity to any genes in the GenBank database beyond the presence of the conserved ANK domains in WD0514 and a coiled coil section between amino acids 1 and 150. Considering that only *Wolbachia* is capable of inducing reproductive incompatibilities in its arthropod hosts and the other related genera are not, it could be expected that *Wolbachia* genes associated with these phenomena would be found only in the *Wolbachia* lineage.

The sequencing of ANK genes in *Wolbachia* strains *w*Mel and *w*Au that infect *Drosophila* and cause different phenotypes has revealed considerable variation between the strains in 10 of the 23 ANK genes. This variation was unexpected given how closely related the *w*Mel and *w*Au strains are (see the cladogram in Fig. 1) and the inability to readily discriminate between them with other molecular markers at the time that

Downloaded from http://jb.asm.org/ on October 13, 2015 by University of Queensland Library

this work was initiated. When the genes are examined in the context of CI expression, five genes are potential candidates for mediating, or at least modulating, CI expression in *Drosophila*. These include WD0514 and its associated operon that is found only in strains that are capable of generating the CI phenotype characteristic of *w*Mel strains. Also, WD0636 and WD0385 are interesting as both of them are disrupted in *w*Au strains, which are incapable of generating CI in *Drosophila*. The disruption of WD0385 in *w*Au by an IS*5* element terminates transcription of this gene in *w*Au, an interesting exception given that all ANK genes are actively expressed in *Wolbachia*, as shown by RT-PCR. Finally, the protein encoded by WD0766 has a different number of repeats in *w*Mel and *w*Au, and in *w*Au it contains a premature stop that eliminates the two transmembrane domains at the C terminus (Fig. 4) of the protein in *w*Mel. Consequently, this sequence variation could modify the affinity and function of the protein not only by affecting the number and location of ANK domains but also by changing the anchoring of the protein in the membrane and therefore its cellular localization in *w*Au compared with *w*Mel. The opposite takes place with the protein resulting from the fusion between WD0754 and WD0753, which adds transmembrane domains to the encoded protein in the non-CI-inducing strain *w*Au. The separation of WD0754 and WD0753 is probably the result of a point mutation in the *w*Mel clade that resulted in the removal of the transmembrane domains from the original ANK protein, with possible subsequent changes affecting its cellular location and folding. Apart from these genes, a number of other genes, including WD0292, WD0294, WD0498, WD0550, and WD0633, display sequence variability between the *w*Au and *w*Mel strains that in some cases results in insertions and deletions of entire ANK motifs in the encoded proteins. Variability of the structure of these proteins and the number of interacting domains might also be associated with phenotypic differences between these strains.

The genetic differences between phage-related ANK genes, such as WD0294, WD0633, and WD0636, seem to have occurred after the phage was inserted into the *Wolbachia* chromosome, and it is improbable that they represent insertion of phages containing different ANK proteins. At least for the P2-like prophage element designated *w*MelWO-B that contains the WD0633 and WD0636 genes, major rearrangements and translocations have taken place, suggesting that this element is inactive (60). WD0294 is in the *w*MelWO-A region that represents a separate insertion in the *Wolbachia* lineage.

Changes in the modular architecture of multiple domain proteins have been shown to affect the folding, function, and specificity of these proteins (30). The reproductive distortions caused by different *Wolbachia* strains in their hosts could be finely tuned by variations in ANK protein architecture that could affect the stability, specificity, and binding properties of *Wolbachia*'s ANK proteins. The variability of phenotypes induced by different *Wolbachia* strains in their hosts is unlikely to be caused by the presence or absence of a "CI gene(s)," a "parthenogenetic gene(s)," or a "male-killing gene(s)," but it is likely to be caused by variation in the binding and affinity properties of the protein(s) responsible. It has recently been shown that deletion of terminal repeats (from the N or C terminus) in ANK proteins can be tolerated to various extents (55). For example, deletion of terminal repeats in the human ANK protein $p16^{INK4a}$ (a CDK inhibitor and tumor suppressor) decreases its unfolding energy, but the internal repeats maintain their structure (61), whereas in the *Drosophila* Notch protein the effects vary depending on the repeats deleted (63). Therefore, changes in stability produced by modification of internal ANK repeats suggest that there is an evolutionary mechanism by which internal deletions minimize the loss of stability and additions or losses at the protein termini are selected.

Because of their modular structure, ANK proteins seem to be highly tolerant to insertions and deletions that affect entire repeats (55), in contrast to changes in the sequences of globular proteins, which are likely to damage the tertiary structure of entire domains (40). Protein repeat variability generally arises from recombination events, intragenic duplication, and deletions (3). It is clear that recombination mechanisms play an important role in shaping *Wolbachia*'s genome (60), and changes in the ANK repeat protein structure through recombination of ANK genes could be a powerful driving force in the evolutionary history of *Wolbachia* by generating novel proteins with possible diverse functions through relatively simple mechanisms (55).

The stability of the variable ANK proteins found in different *Wolbachia* strains, their secretion and interaction in *Drosophila*, and the role that these proteins might have in (i) mediating the establishment of symbiotic associations, (ii) addressing the molecular communication between symbionts and the host cells in which they reside, and (iii) inducing or modulating the reproductive distortions induced by *Wolbachia* remain to be addressed. Unfortunately, functional assignment of *Wolbachia*'s ANK genes cannot be done at present, as no genetic transformation technologies are currently available for this fastidious endosymbiont.

In summary, analysis of the unusually abundant ANK genes in the genome of *Wolbachia* across a number of phenotypically divergent strains has revealed considerable sequence variation in closely related bacterial strains. Correlations with phenotypes in the *Drosophila* host revealed a number of genes that are potential candidates for genes that are associated with the reproductive distortions generated by *Wolbachia*. The variation which we found in the distribution, expression, sequence, ANK domain architecture, and location, as well as the gain or loss of transmembrane domains in almost one-half of the ANK proteins in comparisons of strains that cause CI and strains that are unable to cause CI, are all factors that may affect the specificity, stability, affinity, and, consequently, function of these proteins.

ACKNOWLEDGMENTS

We are very grateful to Manpreet Sidhu and Jenny Gough for technical support. We also thank Jeremy Brownlie, Wolfgang Miller, Steven Sinkins, Kevin Floate, and Elizabeth McGraw for providing constructive comments on the manuscript.

This work was supported by grants from the Australian Research Council.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Andersson, S. G. E., A. Zomorodipour, J. O. Andersson, T. Sicheritz-Ponten, U. C. M. Alsmark, R. M. Podowski, A. K. Naslund, A. S. Eriksson, H. H.**

Downloaded from http://jb.asm.org/ on October 13, 2015 by University of Queensland Library

Winkler, and C. G. Kurland. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature **396:**133–140.
3. **Andrade, M. A., C. Iratxeta-Perez, and C. P. Ponting.** 2001. Protein repeats: structures, functions, and evolution. J Struct. Biol. **134:**117–131.
4. **Beard, C. B., S. L. O'Neill, R. B. Tesh, F. F. Richards, and S. Aksoy.** 1993. Modification of arthropod vector competence via symbiotic bacteria. Parasitol. Today **9:**179–183.
5. **Bordenstein, S. R., F. P. O'Hara, and J. H. Werren.** 2001. *Wolbachia*-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. Nature **409:**707–710.
6. **Bork, P.** 1993. Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally? Proteins **17:**363–374.
7. **Bouchon, D., T. Rigaud, and P. Juchault.** 1998. Evidence for widespread *Wolbachia* infection in isopod crustaceans: molecular identification and host feminization. Proc. R. Soc. Lond. B Biol. Sci. **265:**1081–1090.
8. **Bourtzis, K., S. L. Dobson, H. R. Braig, and S. L. O'Neill.** 1998. Rescuing *Wolbachia* have been overlooked. Nature **391:**852–853.
9. **Braig, H. R., W. Zhou, S. L. Dobson, and S. L. O'Neill.** 1998. Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. J. Bacteriol. **180:**2373–2378.
10. **Breeuwer, J. A., and J. H. Werren.** 1990. Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. Nature **346:**558–560.
11. **Breeuwer, J. A. J., and G. Jacobs.** 1996. *Wolbachia*: intracellular manipulators of mite reproduction. Exp. Appl. Acarol. **20:**421–434.
12. **Caturegli, P., K. M. Asanovich, J. J. Walls, J. S. Bakken, J. E. Madigan, V. L. Popov, and J. S. Dumler.** 2000. *ankA*: an *Ehrlichia phagocytophila* group gene encoding a cytoplasmic protein antigen with ankyrin repeats. Infect. Immun. **68:**5277–5283.
13. **Charlat, S., L. Le Chat, and H. Mercot.** 2003. Characterization of non-cytoplasmic incompatibility inducing *Wolbachia* in two continental African populations of *Drosophila simulans*. Heredity **90:**49–55.
14. **Foster, J., M. Ganatra, I. Kamal, J. Ware, K. Makarova, N. Ivanova, A. Bhattacharyya, V. Kapatral, S. Kumar, J. Posfai, T. Vincze, J. Ingram, L. Moran, A. Lapidus, M. Omelchenko, N. Kyrpides, E. Ghedin, S. Wang, E. Goltsman, V. Joukov, O. Ostrovskaya, K. Tsukerman, M. Mazur, D. Comb, E. Koonin, and B. Slatko.** 2005. The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. PLoS Biol. **3:**e121.
15. **Giordano, R., S. L. O'Neill, and H. M. Robertson.** 1995. *Wolbachia* infections and the expression of cytoplasmic incompatibility in *Drosophila sechellia* and *D. mauritiana*. Genetics **140:**1307–1317.
16. **Hoffmann, A. A.** 1988. Partial cytoplasmic incompatibility between two Australian populations of *Drosophila melanogaster*. Entomol. Exp. Appl. **48:**61–67.
17. **Hoffmann, A. A., D. Clancy, and J. Duncan.** 1996. A naturally-occurring *Wolbachia* infection in *Drosophila simulans* that does not cause cytoplasmic incompatibility. Heredity **76:**1–8.
18. **Hoffmann, A. A., and M. Turelli.** 1988. Unidirectional incompatibility in *Drosophila simulans*: inheritance, geographic variation and fitness effects. Genetics **119:**435–444.
19. **Hoffmann, A. A., M. Turelli, and G. M. Simmons.** 1986. Unidirectional incompatibility between populations of *Drosophila simulans*. Evolution **40:**692–701.
20. **Holmes, D. S., and J. Bonner.** 1973. Preparation, molecular weight, base composition, and secondary structure of giant nuclear ribonucleic acid. Biochemistry **12:**2330–2338.
21. **Jeyaprakash, A., and M. A. Hoy.** 2000. Long PCR improves *Wolbachia* DNA amplification: *wsp* sequences found in 76% of sixty-three arthropod species. Insect Mol. Biol **9:**393–405.
22. **Laven, H.** 1967. Eradication of *Culex pipiens fatigans* through cytoplasmic incompatibility. Nature **216:**383–384.
23. **Letunic, I., R. R. Copley, S. Schmidt, F. D. Ciccarelli, T. Doerks, J. Schultz, C. P. Ponting, and P. Bork.** 2004. SMART 4.0: towards genomic data integration. Nucleic Acids Res. **32(Database issue):**D142—D144.
24. **McGraw, E. A., D. J. Merritt, J. N. Droller, and S. L. O'Neill.** 2002. *Wolbachia* density and virulence attenuation after transfer into a novel host. Proc. Natl. Acad. Sci. USA **99:**2918–2923.
25. **Merçot, H., B. Llorente, M. Jacques, A. Atlan, and C. Montchamp-Moreau.** 1995. Variability within the Seychelles cytoplasmic incompatibility system in *Drosophila simulans*. Genetics **141:**1015–1023.
26. **Merçot, H., and D. Poinsot.** 1998. . . .and discovered on Mount Kilimanjaro. Nature **391:**853.
27. **Michaely, P., and V. Bennett.** 1992. The ANK repeat: a ubiquitous motif involved in macromolecular recognition. Trends Cell Biol. **2:**127–129.
28. **Min, K., and S. Benzer.** 1997. *Wolbachia*, normally a symbiont of *Drosophila*, can be virulent, causing degeneration and early death. Proc. Natl. Acad. Sci. USA **94:**10792–10796.
29. **Mosavi, L. K., T. J. Cammett, D. C. Desrosiers, and Z. Y. Peng.** 2004. The ankyrin repeat as molecular architecture for protein recognition. Protein Sci. **13:**1435–1448.
30. **Mosavi, L. K., D. L. Minor, Jr., and Z. Y. Peng.** 2002. Consensus-derived structural determinants of the ankyrin repeat motif. Proc. Natl. Acad. Sci. USA **99:**16029–16034.
31. **Oh, H. W., M. G. Kim, S. W. Shin, K. S. Bae, Y. J. Ahn, and H. Y. Park.** 2000. Ultrastructural and molecular identification of a *Wolbachia* endosymbiont in a spider, *Nephila clavata*. Insect Mol. Biol. **9:**539–543.
32. **O'Neill, S. L., R. Giordano, A. M. Colbert, T. L. Karr, and H. M. Robertson.** 1992. 16S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects. Proc. Natl. Acad. Sci. USA **89:**2699–2702.
33. **O'Neill, S. L., A. A. Hoffmann, and J. H. Werren.** 1997. Influential passengers: inherited microorganisms and arthropod reproduction. Oxford University Press, Oxford, United Kingdom.
34. **O'Neill, S. L., and T. L. Karr.** 1990. Bidirectional incompatibility between conspecific populations of *Drosophila simulans*. Nature **348:**178–180.
35. **Poinsot, D., K. Bourtzis, G. Markakis, C. Savakis, and H. Mercot.** 1998. *Wolbachia* transfer from *Drosophila melanogaster* into *D. simulans*: host effect and cytoplasmic incompatibility relationships. Genetics **150:**227–237.
36. **Riegler, M., S. Charlat, C. Stauffer, and H. Mercot.** 2004. *Wolbachia* transfer from *Rhagoletis cerasi* to *Drosophila simulans*: investigating the outcomes of host-symbiont coevolution. Appl. Environ. Microbiol. **70:**273–279.
37. **Rousset, F., and M. Solignac.** 1995. Evolution of single and double *Wolbachia* symbioses during speciation in the *Drosophila simulans* complex. Proc. Natl. Acad. Sci. USA **92:**6389–6393.
38. **Rousset, F., D. Vautrin, and M. Solignac.** 1992. Molecular identification of *Wolbachia*, the agent of cytoplasmic incompatibility in *Drosophila simulans*, and variability in relation with host mitochondrial types. Proc. R. Soc. Lond. B Biol. Sci. **247:**163–168.
39. **Rowley, S. M., R. J. Raven, and E. A. McGraw.** 2004. *Wolbachia pipientis* in Australian spiders. Curr. Microbiol. **49:**208–214.
40. **Sagermann, M., W. A. Baase, and B. W. Matthews.** 1999. Structural characterization of an engineered tandem repeat contrasts the importance of context and sequence in protein folding. Proc. Natl. Acad. Sci. USA **96:**6078–6083.
41. **Salzberg, S. L., J. C. Hotopp, A. L. Delcher, M. Pop, D. R. Smith, M. B. Eisen, and W. C. Nelson.** 2005. Serendipitous discovery of *Wolbachia* genomes in multiple *Drosophila* species. Genome Biol. **6:**R23.
42. **Schultz, J., F. Milpetz, P. Bork, and C. P. Ponting.** 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc. Natl. Acad. Sci. USA **95:**5857–5864.
43. **Sedgwick, S. G., and S. J. Smerdon.** 1999. The ankyrin repeat: a diversity of interactions on a common structural framework. Trends Biochem. Sci. **24:**311–316.
44. **Sharp, P. M., and W. H. Li.** 1987. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. **15:**1281–1295.
45. **Sinkins, S. P., and S. L. O'Neill.** 2000. *Wolbachia* as a vehicle to modify insect populations, p. 271–288. *In* A. M. Handler and A. A. James (ed.), Insect transgenesis: methods and applications. CRC Press, Boca Raton, Fla.
46. **Solignac, M., D. Vautrin, and F. Rousset.** 1994. Widespread occurrence of the proteobacteria *Wolbachia* and partial cytoplasmic incompatibility in *Drosophila melanogaster*. C. R. Acad. Sci. Paris Sci. Ser. III Vie/Life Sci. **317:**461–470.
47. **Stouthamer, R., J. A. Breeuwer, and G. D. Hurst.** 1999. *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. Annu. Rev. Microbiol. **53:**71–102.
48. **Stouthamer, R., J. A. Breeuwer, R. F. Luck, and J. H. Werren.** 1993. Molecular identification of microorganisms associated with parthenogenesis. Nature **361:**66–68.
49. **Sun, L. V., M. Riegler, and S. L. O'Neill.** 2003. Development of a physical and genetic map of the virulent *Wolbachia* strain wMelPop. J. Bacteriol **185:**7077–7084.
50. **Swofford, D. L.** 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. S. Sinauer Associates, Sunderland, Mass.
51. **Taylor, M. J., and A. Hoerauf.** 1999. *Wolbachia* bacteria of filarial nematodes. Parasitol. Today **15:**437–442.
52. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The CLUSTAL_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25:**4876–4882.
53. **Tram, U., P. M. Ferree, and W. Sullivan.** 2003. Identification of *Wolbachia*-host interacting factors through cytological analysis. Microbes Infect. **5:**999–1011.
54. **Tram, U., and W. Sullivan.** 2002. Role of delayed nuclear envelope breakdown and mitosis in *Wolbachia*-induced cytoplasmic incompatibility. Science **296:**1124–1126.
55. **Tripp, K. W., and D. Barrick.** 2004. The tolerance of a modular protein to duplication and deletion of internal repeats. J. Mol. Biol. **344:**169–178.
56. **Turelli, M., and A. A. Hoffmann.** 1995. Cytoplasmic incompatibility in *Drosophila simulans*: dynamics and parameter estimates from natural populations. Genetics **140:**1319–1338.
57. **Turelli, M., and A. A. Hoffmann.** 1999. Microbe-induced cytoplasmic incom-

Downloaded from http://jb.asm.org/ on October 13, 2015 by University of Queensland Library

patibility as a mechanism for introducing transgenes into arthropod populations. Insect Mol. Biol **8:**243–255.
58. **Werren, J. H.** 1997. Biology of *Wolbachia*. Annu. Rev. Entomol. **42:**587–609.
59. **Werren, J. H., and D. M. Windsor.** 2000. *Wolbachia* infection frequencies in insects: evidence of a global equilibrium? Proc. R. Soc. Lond. B Biol. Sci. **267:**1277–1285.
60. **Wu, M., L. V. Sun, J. Vamathevan, M. Riegler, R. Deboy, J. C. Brownlie, E. A. McGraw, W. Martin, C. Esser, N. Ahmadinejad, C. Wiegand, R. Madupu, M. J. Beanan, L. M. Brinkac, S. C. Daugherty, A. S. Durkin, J. F. Kolonay, W. C. Nelson, Y. Mohamoud, P. Lee, K. Berry, M. B. Young, T. Utterback, J. Weidman, W. C. Nierman, I. T. Paulsen, K. E. Nelson, H. Tettelin, S. L. O'Neill, and J. A. Eisen.** 2004. Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. PloS Biol. **2:**327–341.
61. **Zhang, B., and Z. Peng.** 2000. A minimum folding unit in the ankyrin repeat protein p16INK4. J. Mol. Biol. **299:**1121–1132.
62. **Zhou, W., F. Rousset, and S. L. O'Neill.** 1998. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. Proc. R. Soc. Lond. B Biol. Sci. **265:**509–515.
63. **Zweifel, M. E., and D. Barrick.** 2001. Studies of the ankyrin repeats of the *Drosophila melanogaster* Notch receptor. 1. Solution conformational and hydrodynamic properties. Biochemistry **40:**14344–14356.

Downloaded from http://jb.asm.org/ on October 13, 2015 by University of Queensland Library